

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-024981

(43)Date of publication of application : 29.01.1999

(51)Int.Cl.

G06F 12/00

(21)Application number : 09-178534

(71)Applicant : NEC CORP

(22)Date of filing : 03.07.1997

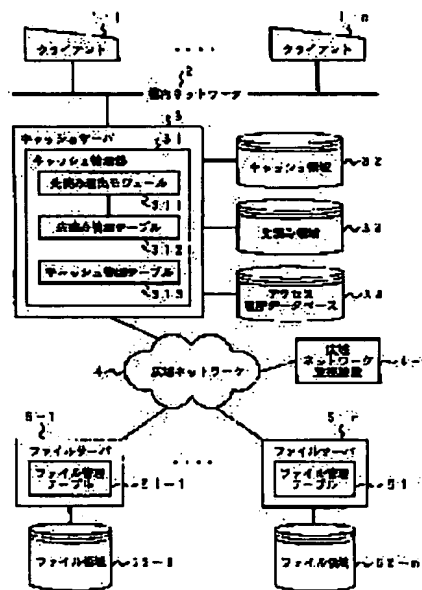
(72)Inventor : UEHARA OSAMU

## (54) WIDE AREA DISTRIBUTED FILE SYSTEM

## (57)Abstract:

**PROBLEM TO BE SOLVED:** To provide a wide area distributed file system where a cache hit ratio is improved and file access is accelerated.

**SOLUTION:** A cache server 3 has an access history data base 34 for recording file access. A prefetch selection module 311 analyzes a file together with its updating frequency whose access frequency is higher than an access history data base and determines a file which is to be looked ahead. High speed file access is provided independent of the state of a wide area network file by looking ahead the file whose access frequency is high in a time zone when the wide band network is not congested and transferring the file which is looked ahead against a request from a client.



## LEGAL STATUS

[Date of request for examination] 03.07.1997

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 3117003

[Date of registration] 06.10.2000

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right] 06.10.2004

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-24981

(43) 公開日 平成11年(1999) 1月29日

(51) Int.Cl.  
G 0 6 F 12/00

識別記号  
5 4 5

F I  
G 0 6 F 12/00

5 4 5 Z

審査請求 有 請求項の数 4 O L (全 8 頁)

(21) 出願番号 特願平9-178534

(22) 出願日 平成9年(1997) 7月3日

(71) 出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72) 発明者 上原 修

東京都港区芝五丁目7番1号 日本電気株式会社内

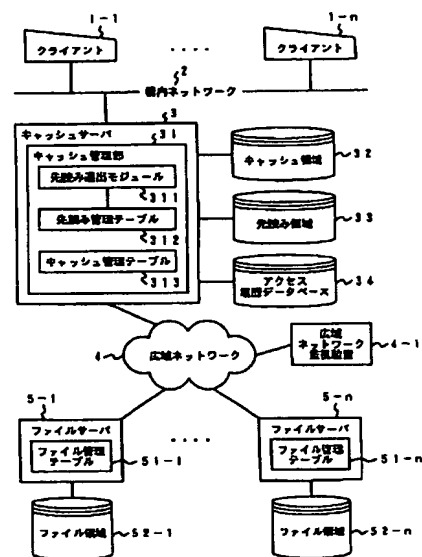
(74) 代理人 弁理士 後藤 洋介 (外2名)

(54) 【発明の名称】 広域分散ファイルシステム

(57) 【要約】

【課題】 キャッシュヒット率を向上しファイルアクセスを高速化する広域分散ファイルシステムを提供することである。

【解決手段】 キャッシュサーバ3はファイルアクセスを記録するアクセス履歴データベース34をもつ。先読み選出モジュール311はアクセス履歴データベースよりアクセス頻度の高いファイルとその更新間隔を分析し、先読みすべきファイルを決定する。広域ネットワークがすいている時間帯にアクセス頻度の高いファイルを先読みし、クライアントからの要求に対し先読みしたファイルを渡すことにより広域ネットワークの状態に関わらず高速なファイルアクセスを提供する。



【特許請求の範囲】

【請求項1】 複数のファイルを保管する複数のファイルサーバと、該ファイルサーバに保管されたファイルにアクセスするクライアントと、前記ファイルを保管するキャッシュサーバが広域ネットワークに接続されている場合において、前記キャッシュサーバは、キャッシュ管理部、通常のファイルキャッシュに使われるキャッシュ領域、先読みを行ったファイルをキャッシュする先読み領域、及びクライアントからのアクセス要求の履歴を記録するアクセス履歴データベースを具備して構成され、定期的に前記アクセス履歴データベース内のファイルアクセスの履歴を分析し、アクセス頻度の高いファイルを割り出し、定期的に前記キャッシュサーバへ先読みすることにより、ファイルアクセスを高速化することを特徴とする広域分散ファイルシステム。

【請求項2】 前記キャッシュ管理部は、定期的にアクセス履歴データベースを分析し先読みするファイルと読み込み間隔を決定する先読み選出モジュールと、決定したファイル名と間隔を記録する先読み管理テーブルと、前記キャッシュ領域を管理するキャッシュ管理テーブルを具備して構成されていることを特徴とする請求項1記載の広域分散ファイルシステム。

【請求項3】 前記広域ネットワークには、ネットワークの混雑や障害情報を管理するネットワーク監視装置が設置されていることを特徴とする請求項1又は2記載の広域分散ファイルシステム。

【請求項4】 前記複数のファイルサーバは、ファイルを格納するファイル領域とそれらを管理するファイル管理テーブルを具備していることを特徴とする請求項1乃至3のいずれか一つに記載の広域分散ファイルシステム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、ファイルを保管するファイルサーバとファイルを参照するクライアントとが広域ネットワークで接続された分散ファイルシステムに関し、特に使用頻度の高いファイルをキャッシングする分散ファイルシステムに関するものである。

【0002】

【従来の技術】 多くの分散ファイルシステムでは、ファイルサーバが保管するファイルをクライアントにキャッシングすることで、ファイルアクセスを高速化している。すなわち、ファイルサーバが保管するファイルのコピーをクライアント側で保持しておくことにより、次回そのファイルに対するアクセス要求が発生した場合には、ファイルサーバへアクセスすることなく自クライアント内で処理することができる。近年は広域ネットワークを利用して分散ファイルシステムを構築する事もあり、この場合には構内ネットワークと広域ネットワークの間にキャッシュサーバと呼ばれるキャッシングを専門と

したサーバが設置され、構内ネットワークに接続しているクライアントはキャッシュサーバを通じてファイルサーバからのファイルを取得する。ファイルアクセス性能に影響を与える要因にはファイルサーバの性能、ネットワークのファイル転送速度、キャッシュのヒット率などがある。

【0003】 キャッシュのヒット率を上げる従来技術として、特開平7-93205号公報に開示されている技術やCoda分散ファイルシステムがある。これらの技術ではユーザやクライアント毎にファイルの優先度を設定する。ファイルアクセス統計をもとに頻繁にアクセスするファイルを選び、高い優先度に設定し高い確率でキャッシュ領域に残すことで、キャッシュのヒット率を向上させている。

【0004】

【発明が解決しようとする課題】 しかしながら上記従来技術では、広域ネットワークは構内ネットワークに比べ転送速度が遅く、また利用者が多いため時間帯やタイミングによって混雑状況が大きく変わり、また、広域ネットワークの状態によってはファイルサーバに接続できないこともあり得る。このため、ファイルがキャッシュ領域になかったり、ファイルサーバ側で更新されていた場合にはファイルサーバからファイルを取得するわけだが、この時広域ネットワークを介してファイル転送を行うので、ファイルサーバからキャッシュサーバへの転送に時間がかかり、場合によっては転送ができないことも生ずる。

【0005】 又、ファイルの優先度をユーザやクライアント毎に設定しているため、部門が同じユーザは同じファイルを頻繁に参照すると思われるが、現在のキャッシュ方式では部門全体のキャッシュヒット率を上げることはできない。

【0006】 本発明の目的は、キャッシュのヒット率を向上し、ファイルアクセスを高速化する広域分散ファイルシステムを提供することである。

【0007】

【課題を解決するための手段】 本発明によれば、複数のファイルを保管する複数のファイルサーバと、該ファイルサーバに保管されたファイルにアクセスするクライアントと、前記ファイルを保管するキャッシュサーバが広域ネットワークに接続されている場合において、前記キャッシュサーバは、キャッシュ管理部、通常のファイルキャッシュに使われるキャッシュ領域、先読みを行ったファイルをキャッシュする先読み領域、及びクライアントからのアクセス要求の履歴を記録するアクセス履歴データベースを具備して構成され、定期的に前記アクセス履歴データベース内のファイルアクセスの履歴を分析し、アクセス頻度の高いファイルを割り出し、定期的に前記キャッシュサーバへ先読みすることにより、ファイルアクセスを高速化することを特徴とする広域分散フ

イルシステムが得られる。

【0008】さらに、本発明によれば、前記キャッシュ管理部は、定期的にアクセス履歴データベースを分析し先読みするファイルと読み込み間隔を決定する先読み選出モジュールと、決定したファイル名と間隔を記録する先読み管理テーブルと、前記キャッシュ領域を管理するキャッシュ管理テーブルを具備して構成されていることを特徴とする広域分散ファイルシステムが得られる。

【0009】さらに、本発明によれば、前記広域ネットワークには、ネットワークの混雑や障害情報を管理するネットワーク監視装置が設置されていることを特徴とする広域分散ファイルシステムが得られる。

【0010】さらに、本発明によれば、前記複数のファイルサーバは、ファイルを格納するファイル領域とそれらを管理するファイル管理テーブルを具備していることを特徴とする広域分散ファイルシステムが得られる。

【0011】

【作用】キャッシュサーバはクライアントのファイルアクセスの記録をアクセス履歴データベースに記録している。先読み選出モジュールがアクセス履歴データベースを分析することで、そのキャッシュサーバを使用している部門が頻繁にアクセスしているファイルと更新間隔を分析することができる。

【0012】又、キャッシュサーバは、先読み選出モジュールが決定した先読みファイルを、広域ネットワーク監視装置が指定した時間に、先読み領域に読み込む。キャッシュサーバはクライアントが要求するファイルが先読み領域かキャッシュ領域のいずれかにあればそこからファイルを取得する。通常のキャッシュに加えあらかじめ参照頻度の高いファイルを先読みすることによりヒット率が向上し、ファイルアクセスが高速化する。

【0013】又、部門という業務単位毎にキャッシュサーバを設置することにより、その業務に関連するファイルがキャッシュサーバに先読み、キャッシュされ、ヒット率が向上する。

【0014】

【発明の実施の形態】以下、本発明の一実施の形態について図面を参照して詳細に説明する。図1に示すように、本発明に係る広域分散ファイルシステムは、複数のファイルを保管する複数のファイルサーバ5-1～5-nと、構内ネットワーク2に接続され、ファイルをキャッシュするキャッシュサーバ3と、キャッシュサーバ3に接続されている広域ネットワーク4と、ファイルサーバに保管されたファイルにキャッシュサーバ3を介してアクセスするクライアント1-1～1-nを具備して構成される。

【0015】キャッシュサーバ3は、キャッシュ管理部31、通常のファイルキャッシュに使われるキャッシュ領域32、先読みを行ったファイルをキャッシュする先読み領域33、及びクライアントからのアクセス要求の

履歴を記録するアクセス履歴データベース34を具備して構成される。

【0016】キャッシュ管理部31は、定期的にアクセス履歴データベースを分析し先読みするファイルと読み込み間隔を決定する先読み選出モジュール311と、決定したファイル名と間隔を記録する先読み管理テーブル312、キャッシュ領域を管理するキャッシュ管理テーブル313を具備して構成される。

【0017】また、広域ネットワーク4には、ネットワークの混雑や障害情報を管理するネットワーク監視装置4-1が設置されている。ファイルサーバは、ファイルを格納するファイル領域52-nとそれを管理するファイル管理テーブル51-nを持つ。

【0018】次に本実施の形態の動作について、図1～図4を参照して詳細に説明する。クライアント1-1からキャッシュサーバ3へファイルアクセス要求があると（ステップS6）、キャッシュ管理部31は対象ファイルがキャッシュサーバ内に存在するか否かをまず先読み管理テーブル312で、次にキャッシュ管理テーブル313で確認する（ステップS61、ステップS62）。

【0019】もしファイルが存在するならば、ファイルサーバに接続し（ステップS610、ステップS621）、その先読み領域のファイル、キャッシュ領域のファイルが最新のものであるか否かをファイルサーバに問い合わせる（ステップS611、ステップS622）。その結果先読み領域またはキャッシュ領域にあるファイルが最新であれば、そのファイルをクライアントに渡す（ステップS618、ステップS67）。

【0020】ファイルサーバへの接続が失敗したときは、クライアントにワーニングを通知し（ステップS617）、無条件に先読み領域またはキャッシュ領域にあるファイルをクライアントに渡すファイルとする（ステップS618、ステップS67）。これにより広域ネットワークの状態によってファイルサーバに接続できなかったとしても、クライアントにファイルを渡すことができる。

【0021】ファイルが最新のものでないならば、ファイルサーバから最新のものを取得し、現在あるファイルと置き換える。対象ファイルが先読み領域の場合、新しいファイルのサイズが今あるファイルよりも大きい場合には（ステップS613）、サイズを確保するために優先度の低いファイルから先読み領域から追い出してサイズを確保して、ファイルを置き換え、先読み管理テーブル312の更新処理を行い（ステップS614、ステップS615）、追い出したファイルをキャッシュ領域に移す（ステップS616）、そうでないなら現在のファイルを新しいものに置き換える（ステップS619、ステップS620）。対象ファイルがキャッシュ領域の場合、または、先読み領域からキャッシュ領域に追い出されたファイルが存在する場合、キャッシュ管理部31

はキャッシュ領域のサイズに合わせて、キャッシュ更新処理を行う（ステップS 6 5、ステップS 6 6）。処理手順は通常のキャッシュ更新処理と同様なのでここでは述べない。

【0022】ファイルが先読み領域またはキャッシュ領域に存在しない時は、キャッシュサーバはファイルサーバに接続し（ステップS 6 3）、ファイルを取得する（ステップS 6 4）。取得したファイルはキャッシュ領域にキャッシュする（ステップS 6 5、ステップS 6 6）。ファイルサーバとの接続に失敗したときは、クライアントに渡すファイルを取得できないため、エラーをクライアントに通知する（ステップS 6 8）。クライアントからのファイル取得要求はすべてアクセス履歴データベース34に記録する（ステップS 6 9）。

【0023】キャッシュ管理部31は定期的に先読み選出モジュール311を読み出し、先読みすべきファイルを選出し、先読み管理テーブルを作成する（ステップS 7）。先読みすべきファイルは、アクセス履歴データベースより参照回数の多い物から選出する（ステップS 7 1）。選出したファイルに関し、アクセス履歴データベース34から更新の履歴を割り出し分析することによりファイルの更新間隔を推測する。推測した更新間隔を先読み間隔として設定する（ステップS 7 2）。

【0024】選出したファイルに関し、広域ネットワーク監視装置4-1からネットワークがあいている時間を割り出し更新のタイミングとする（ステップS 7 3）。以上の結果より先読み管理テーブル312を更新する（ステップS 7 4）。キャッシュ管理部31は定期的に先読み管理テーブル312を参照し、先読み間隔と最終読み込み日時から先読みが必要なファイルを割り出し、先読み処理を行う（ステップS 8）。処理手順はクライアントが先読み領域内のファイルを読み出すときに行う処理と同様である（ステップS 6 10からの処理と同様）。

【0025】

【実施例】次に、本発明の実施例について図1～図4に加え、図5及び図6を参照して説明する。図5及び図6は実装した各管理テーブルの項目とその状態の一実施例を示した図である。

【0026】ファイル管理テーブル51-1～51-nでは、管理するファイルのファイル名とサイズ、更新日時を管理する。アクセス履歴データベース34は、キャッシュしたファイルの存在するファイルサーバ名とそのファイル名、ファイルのサイズ、ファイルサーバに記録されていたそのファイルの更新日時、クライアントのアクセス日時を記録する。クライアントのアクセス要求が生じる度に（ファイルがキャッシュ領域、先読み領域、ファイルサーバのいずれかにあろうとも）アクセス履歴データベース34に記録される。

【0027】先読み管理テーブル312は、先読み領域

内のファイル名とそのファイルの存在するファイルサーバ名、ファイルのサイズ、先読みしたときの更新日時、最終先読み日時、先読みモジュールが決定した先読み間隔を管理する。先読み管理テーブル312は先読み選出モジュールがアクセス履歴データベース34を参照して、先読みすべきファイルとその先読み間隔を決定することにより作成される。また、更新日時はクライアントからの要求によるファイルサーバからの読み込み（図3のステップS 6 15、ステップS 6 20）によっても更新される。

【0028】キャッシュ管理テーブル313は、キャッシュ領域内のファイル名とそのファイルの存在するファイルサーバ名、ファイルのサイズ、更新日時を管理する。実装例の説明のためにキャッシュ領域と先読み領域のサイズ、各テーブルの項目に仮の値を設定しておく。

【0029】次に、クライアントからのファイル要求について図5を参照して説明する。クライアントからファイルf1-1の読み出し要求があると、キャッシュ管理部31はファイルf1-1が先読み領域にあるかどうかを先読み管理テーブル312より確認する（ステップS 6 1）。ファイルf1-1は先読み管理テーブル312にあるため、このファイルが最新であるかどうかをファイルサーバ51のファイル管理テーブル51-1より確認する（ステップS 6 10、ステップS 6 11）。このときファイルサーバ51に接続できないときは（ステップS 6 10で失敗）、クライアントに警告を出し（ステップS 6 17）、先読み領域にあるファイルf1-1をクライアントに渡す（ステップS 6 7）。ファイルサーバ51にファイルと同じ更新日時であれば先読み領域にあるファイルf1-1をクライアントに渡すが、そうでないのでファイルサーバ51からファイルf1-1を取得する（ステップS 6 12）。読み込んだ新しいファイルf1-1はサイズが増えているので先読み領域には入りきれない。そこで先読み管理テーブル312より、優先度の低いファイルf2-1を追い出し、キャッシュ領域に移動する（ステップS 6 16）。キャッシュ領域は通常のキャッシュ管理手順でファイルf2-1を処理する（ステップS 6 5）。もしFIFO式の管理ならばファイルf1-3を追い出しファイルf2-1を挿入する。アクセス履歴データベース34にはファイルf1-1の情報とアクセス日時が記録される（ステップS 6 9）。各管理テーブルは変更され図6となる。

【0030】次に先読みファイル選出処理について図6を参照して説明する。先読み選出モジュール311はアクセス履歴データベース34より参照の多いファイルf1-1、f2-2、f1-2を選び出す（ステップS 7 1）。次にそれぞれのファイルについて過去の更新日時を比較し、ファイルの更新間隔を推測すると、6日、2日、7日となる（ステップS 7 2）。また広域ネットワーク監視装置4-1のデータより、広域ネットワークは

午前1時にもっとも効率よくファイル転送が行えることがわかった（ステップS73）。そこでこれらの情報を元に先読み管理テーブルを更新する（ステップS74）。更新は変更されたところのみ行う。この例であればファイルf1-1は先読み間隔が6日に、最終読み込み日時の時間が午前1時に更新され、ファイルf1-2の項目が新たに追加される。新たに追加された項目（ここではファイルf1-2）については、最終読み込み日時のうち日付を空欄にすることで次に午前1時になったときかならず先読みするようにする。各管理テーブルは更新され図7となる。

【0031】キャッシュ管理部31は定期的先読み管理テーブル312をチェックしている。もし最終読み込み日時と先読み間隔を足した日時が過ぎるか、最終読み込み日時の日付が空欄のもので指定する時間になったものがあつた場合、先読み処理を行う。ここではファイルf1-2に関して、午前1時に先読みを行う。手順はクライアントが先読み領域内のファイルを読み出すときに行う処理と同様である（ステップS610からの処理と同様）。

【0032】

【発明の効果】本発明によれば、参照頻度の高いファイ

ルをあらかじめキャッシュサーバに先読みしているため、広域ネットワークの転送速度が遅く、場合によってファイルサーバに接続できないとしても、クライアントに迅速にファイルを転送することができる。

【0033】又、本発明によれば、部門ごとにキャッシュサーバを設置するため、部門単位でのキャッシュヒット率を上げることができる。

【図面の簡単な説明】

【図1】本発明に係る広域分散ファイルシステムの一実施の形態を表す構成図である。

【図2】クライアントからのアクセス処理要求のフローチャートである。

【図3】図2のクライアントからのアクセス処理要求のフローチャートの続きを示したフローチャートである。

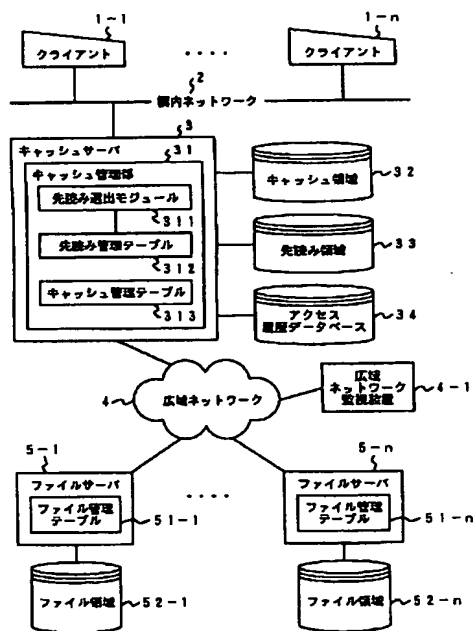
【図4】先読みファイル選出処理および先読み処理のフローチャートである。

【図5】本実施例を説明するために用いた各管理テーブル値である。

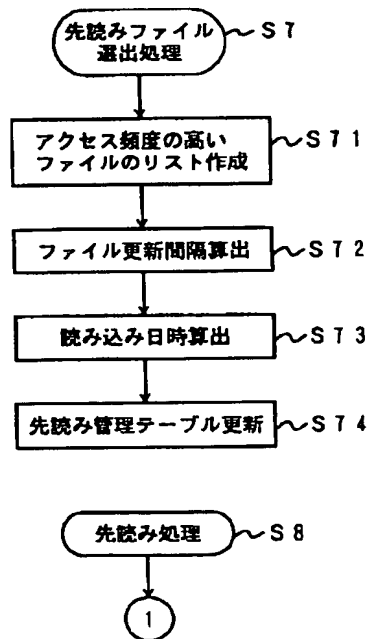
【図6】本実施例を説明するために用いた各管理テーブル値である。

【図7】本実施例を説明するために用いた各管理テーブル値である。

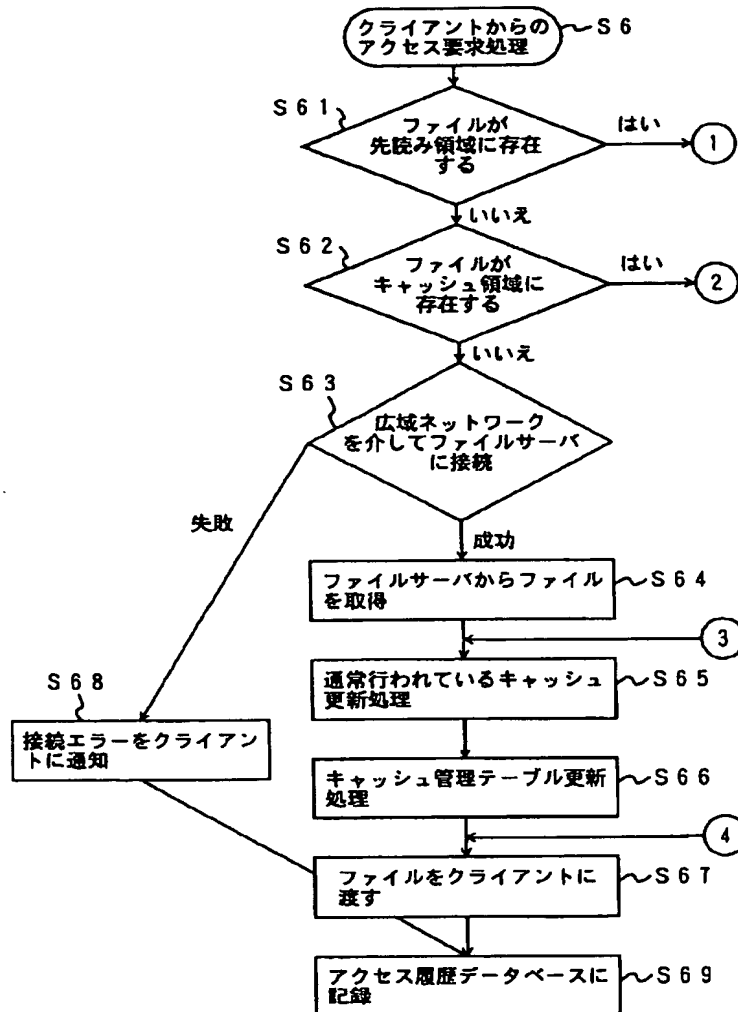
【図1】



【図4】



【図2】

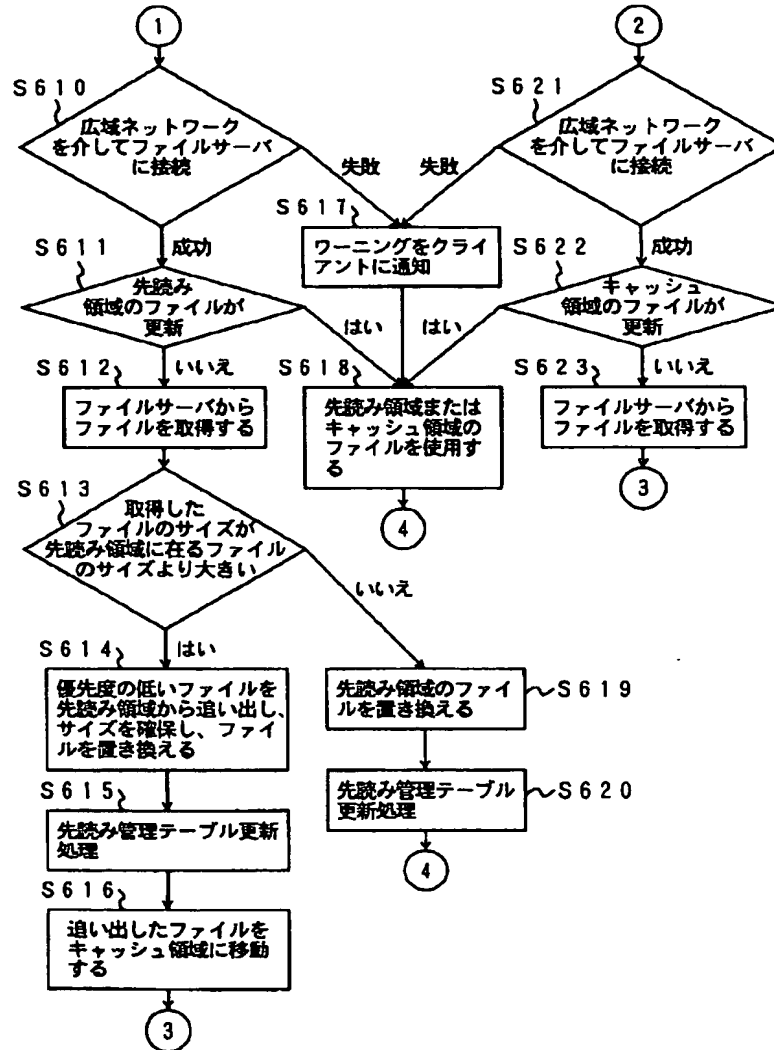


【図7】

先読み管理テーブル (サイズ500,000) (312)

ファイル サーバ名	ファイル名	サイズ	更新日時	最終読み込み 日時	先読み開始
S1-1	F1-1	200,000	980130 120000	980131 010000	000000 000000
S1-2	F2-2	150,000	980129 120000	980130 010000	000002 000000
S1-1	F1-2			010000	000007 000000

【図3】





【図5】

ファイル管理テーブル (51-1)			ファイル管理テーブル (51-2)		
ファイル名	サイズ	更新日時	ファイル名	サイズ	更新日時
f1-1	800,000	880130 120000	f2-1	70,000	880129 120000
f1-2	40,000	880124 120000	f2-2	150,000	880129 120000
f1-3	500,000	880124 120000	f2-3	600,000	880129 120000

キャッシュ管理テーブル (サイズ1,000,000) (313)

ファイル サーバ名	ファイル名	サイズ	更新日時
s1-1	f1-3	500,000	880124 120000
s1-2	f2-3	400,000	880129 120000
s1-3	f1-2	40,000	880124 120000

先読み管理テーブル (サイズ500,000) (312)

ファイル サーバ名	ファイル名	サイズ	更新日時	最終読み込み 日時	先読み回数
s1-1	f1-1	200,000	880130 120000	880129 010000	080007 000000
s1-2	f2-2	150,000	880129 120000	880129 010000	080002 000000
s1-3	f2-1	70,000	880129 120000	880129 010000	080002 000000

【図6】

ファイル管理テーブル (51-1)			ファイル管理テーブル (51-2)		
ファイル名	サイズ	更新日時	ファイル名	サイズ	更新日時
f1-1	850,000	880130 120000	f2-1	70,000	880129 120000
f1-2	40,000	880124 120000	f2-2	150,000	880129 120000
f1-3	500,000	880124 120000	f2-3	400,000	880129 120000

キャッシュ管理テーブル (サイズ1,000,000) (313)

ファイル サーバ名	ファイル名	サイズ	更新日時
s1-1	f1-3	500,000	880124 120000
s1-2	f2-3	400,000	880129 120000
s1-3	f2-1	70,000	880129 120000

先読み管理テーブル (サイズ500,000) (312)

ファイル サーバ名	ファイル名	サイズ	更新日時	最終読み込み 日時	先読み回数
s1-1	f1-1	200,000	880130 120000	880131 141400	080007 000000
s1-2	f2-2	150,000	880129 120000	880130 010000	080002 000000

アクセス履歴データベース (34)

ファイル サーバ名	ファイル名	サイズ	更新日時	アクセス日時
s1-1	f1-2	80,000	880117 120000	880129 120000
s1-2	f2-4	150,000	880127 120000	880129 120000
s1-3	f1-1	200,000	880124 120000	880129 140000
s1-4	f2-2	150,000	880129 120000	880130 140000
s1-1	f1-2	80,000	880124 120000	880130 140000
s1-3	f1-1	200,000	880130 120000	880130 141400
s1-4	f1-1	200,000	880130 120000	880130 150000
s1-1	f1-1	200,000	880130 120000	880130 180000
s1-2	f2-2	150,000	880129 120000	880130 181600